WHAT IS CLAIMED IS:

1. A method for performing morphology analysis of a natural language document, comprising the steps of:

inputting the natural language document as an input text,

tokenize said input text, thereby producing a token stream,

checking for each token of said token stream, whether it is a unique token which is occurring for the first time in the token stream or a recurring token which already occurred earlier in the token stream,

marking unique tokens with an identification (ID) and adding to recurring tokens a pointer directed towards the ID which was defined for the respective token when occurring for the first time,

performing a morphological look-up only on the unique tokens thereby producing results of morphological look-up,

storing said results of morphological look-up for the unique tokens together with the ID,

reading the results of morphological look-up for the recurring tokens,

joining all results of morphological look-up, thereby producing a stream of morphological analyses,

outputting said stream of morphological analyses.

2. The method according to claim 1, wherein the step of storing said results of morphological look-up for the unique tokens together with the ID comprises the step of creating a dynamically extending database.

3. The method according to claim 2, wherein said dynamically extending database is a self-extending hash table.

4. The method according to claim 1, wherein the step of tokenizing is performed by a first finite state transducer.

5. The method according to claim 4, wherein the first finite state transducer includes punctuation conventions and higher level lexical information.

6. The method according to claim 4, wherein the step of morphological look-up is performed by a second finite state transducer.

7. A system for performing morphology analysis of a natural language document, comprising:

a tokenizer for tokenizing an input document, thereby producing a token stream,

a pre-processor that checks for each token of said token stream, whether it is a unique token which is occurring for the first time in the token stream or a recurring token which already occurred earlier in the token stream, marks unique tokens with an identification (ID) and adds to recurring tokens a pointer directed towards the ID which was defined for the respective token when occurring for the first time,

a morphological look-up module for performing a morphological look-up only on the unique tokens, thereby producing results of morphological look-up,

memory for storing said results of morphological look-up for the unique tokens together with the ID,

a post-processor that detects tokens carrying said pointer and replaces them by said results of morphological look-up stored in the memory under the respective ID.

8. The system according to claim 7, wherein the tokenizer is a first finite state transducer.

9. The system according to claim 8, wherein the tokenizer includes punctuation conventions and higher level lexical information.

10. The system according to claim 8, wherein the morphological look-up module is a second finite state transducer.

11. The system according to claim 7, wherein the memory is a dynamically extending database.

12. The system according to claim 11, wherein said dynamically extending database is a self-extending hash table.

13. A system for performing morphology analysis of a natural language document, comprising:

a tokenizer for tokenizing an input document, thereby producing a token stream,

a morphological look-up module for performing the morphological look-up, thereby producing results of morphological look-up,

memory for storing said results of morphological look-up for the unique tokens together with the ID,

a control unit for controlling said morphological look-up module and said memory; wherein said control unit:

checks for each token of said token stream, whether it is a unique token which is occurring for the first time in the token stream or a recurring token which already occurred earlier in the token stream, and

marks unique tokens with an identification (ID) and adds to recurring tokens a pointer directed towards the ID which was defined for the respective token when occurring for the first time,

initializes a morphological look-up for the unique tokens and replaces recurring tokens marked with said pointer by results of morphological look-up stored in the memory under the respective ID.

14. The system according to claim 13, wherein the tokenizer is a first finite state transducer.

15. The system according to claim 14, wherein the tokenizer includes punctuation conventions and higher level lexical information.

16. The system according to claim 14, wherein the morphological look-up module is a second finite state transducer.

17. The system according to claim 13, wherein the memory is a dynamically extending database.

18. The system according to claim 17, wherein said dynamically extending database is a self-extending hash table.